EXHIBIT A

Protein Engineering vol.12 no.7 pp.597–604, 1999

scFv multimers of the anti-neuraminidase antibody NC10: length of the linker between V_H and V_L domains dictates precisely the transition between diabodies and triabodies

John L.Atwell[1], Kerry A.Breheney, Lynne J.Lawrence[2], Airlie J.McCoy[3], Alexander A.Kortt and Peter J.Hudson

CSIRO Molecular Science and CRC for Diagnostic Technologies, 343 Royal Parade, Parkville, Victoria, [2]Biomolecular Research Institute, 343 Royal Parade, Parkville, Victoria, Australia 3052 and [3]Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[1]To whom correspondence should be addressed.
E-mail: john.atwell@molsci.csiro.au

Single-chain Fv antibody fragments (scFvs) incorporate a polypeptide linker to tether the V_H and V_L domains together. An scFv molecule with a linker 5–12 residues long cannot fold into a functional Fv domain and instead associates with a second scFv molecule to form a bivalent dimer (diabody). Direct ligation of V_H and V_L domains further restricts association and forces three scFv molecules to associate into a trivalent trimer (triabody). We have defined the effect of linker length on scFv association by constructing a series of scFvs from anti-neuraminidase antibody NC10 in which the linker varied from one to four glycine residues. NC10 scFv molecules containing linkers of three and four residues showed a strong preference for dimer formation (diabodies), whereas a linker length of one or two glycine residues prevented the formation of diabodies and directed scFv association into trimers (triabodies). The data suggest a relatively strict transition from dimer (diabody) to trimer (triabody) upon reduction of the linker length from three to two glycine residues. Modelling studies are consistent with three residues as the minimum linker length compatible with diabody formation. Electron microscope images of complexes formed between the NC10 scFv multimers and an anti-idiotype Fab′ showed that the dimer was bivalent for antigen binding and the trimer was trivalent.

Keywords: antibody/diabody/dimers/single-chain Fvs/shorter linkers/triabody/trimers

Introduction

The antigen binding domain of an antibody can be expressed in *Escherichia coli* as a single-chain molecule (scFv) in which the V_H and V_L domains of the antibody are joined by a flexible polypeptide linker (Bird *et al.*, 1988; Huston *et al.*, 1988). When the linker is >12 amino acids in length, the V_H and V_L domains dock with each other in the natural Fv orientation to form the antigen combining site (Kortt *et al.*, 1994; Whitlow *et al.*, 1994). The scFv usually shows a monovalent antigen binding affinity similar to the Fab fragment of the parent antibody (Kortt *et al.*, 1994; Zdanov *et al.*, 1994). As the linker is shortened to ≤12 residues, the V_H domain is unable to bind to its attached V_L domain in the Fv orientation. Instead, complementary V_H/V_L pairs from two separate scFv molecules combine to form a bivalent dimer, termed a diabody (Holliger *et al.*, 1993) If the two scFv molecules are identical, two equivalent Fv modules form and the diabody is bivalent but monospecific (Kortt *et al.*, 1997). Bispecific diabodies are made by combination of two different scFv molecules where the V_L of an scFv of one specificity is joined via a shortened linker (<12 residues) to the V_H of an scFv of another specificity and vice versa (Holliger *et al.*, 1993, 1996; Perisic *et al.*, 1994; Atwell *et al.*, 1996). When the linker is removed completely and the C-terminal residue of the V_H domain is linked directly to the N-terminal residue of the V_L domain, three scFv molecules associate to form a trimer, termed a triabody. Triabodies described so far have been either monospecific with three equivalent antigen binding sites (trivalent) (Iliades *et al.*, 1997; Kortt *et al.*, 1997) or non-functional (non-cogante V_H/V_L pair) (Pei *et al.*, 1997). These multivalent molecules can provide high avidity to target antigens and are sufficiently large to avoid the fast clearance through the kidneys observed for scFv monomers and thereby have potential application for tumour imaging and radiotherapy.

Modelling studies based on the X-ray crystal structure of a diabody (five-residue linker scFv) showed that reducing the linker to 'just one or two residues' seriously strained the diabody association (Perisic *et al.*, 1994). The definition of linker length is imprecise because the last residue held in contact within the V_H domain framework is unique for each antibody. For example, the NC10 (anti-neuraminidase) triabody was defined as a linkerless scFv with direct ligation of the terminal residue in V_H (Ser^{H112}) to the N-terminal residue of V_L (Asp^{L1}), whereas the B1–8/NQ11 triabody (Pei *et al.*, 1997) and 11–1G10 triabody (Iliades *et al.*, 1997) included an additional residue, Ser^{H113}, between V_H and V_L domains. Indeed, it was noted that in the B1–8/NQ11 triabody, the orientation of V_H to V_L domains in the Fv modules was distorted and a CDR loop structure was misplaced. A report of a linkerless scFv monospecific 'diabody' library (McGuiness *et al.*, 1996) further confused the field since the molecular mass of these 'diabodies' was not determined and many were likely to be triabodies.

To clarify the effect of linker length on scFv association, we have modelled NC10 diabodies and triabodies, constructed a series of scFvs from NC10 with linkers of one, two, three and four amino acids and determined which linker length changes the preferred conformation of the multimer from a dimer to a trimer. Our results show that linkers of three residues or greater are required for diabody formation whereas two residues or less enable scFvs to associate into active trimers.

Materials and methods

Sequence numbering

Antibody residues were numbered according to Kabat *et al.* (1991) and for NC10 correspond exactly to Malby *et al.* (1993). Residues in the V_H domain of the scFv were superscripted with H and the residue number; for example Ser^{H112} signifies serine in position 112 of the V_H domain. Similarly, residues in

Table I. Oligonucleotide primer pairs used to mutate NC10 scFv–0 to produce scFv–1, –2, –3 and –4, using QuikChange protocol (additional glycine codons are denoted in lower case)

scFv–1
antisense
5′ GGG ACC ACG GTC ACC GTC TCC ggt GAT ATC GAG CTC ACA CAG 3′
sense
5′ CTG TGT GAG CTC GAT ATC acc GGA GAC GGT GAC CGT GGT CCC 3′
scFv–2
antisense
5′ GGG ACC ACG GTC ACC GTC TCC ggt ggt GAT ATC GAG CTC ACA CAG 3′
sense
5′ CTG TGT GAG CTC GAT ATC acc acc GGA GAC GGT GAC CGT GGT CCC 3′
scFv–3
antisense
5′ GGG ACC ACG GTC ACC GTC TCC ggt ggt ggt GAT ATC GAG CTC ACA CAG 3′
sense
5′ CTG TGT GAG CTC GAT ATC acc acc acc GGA GAC GGT GAC CGT GGT CCC 3′
scFv–4
antisense
5′ GGG ACC ACG GTC ACC GTC TCC ggt ggt ggt ggt GAT ATC GAG CTC ACA CAG 3′
sense
5′ CTG TGT GAG CTC GAT ATC acc acc acc acc GGA GAC GGT GAC CGT GGT CCC 3′

the V_L domain of the scFv were superscripted L and the residue number.

Molecular modelling

Computer-generated models of NC10 scFv diabodies and triabodies were constructed using Fv modules that corresponded to the coordinates of the NC10 Fv domain in PDB entry 1NMB (Malby *et al.*, 1994, 1998). Fv modules were manipulated as rigid bodies with the O molecular graphics package (Jones *et al.*, 1991) such that the diabody structure and gross alignment of linkers corresponded to the crystal structure described by Perisic *et al.* (1994) and the triabodies corresponded to the model described by Kortt *et al.* (1997). Thus, the diabodies comprised two Fv modules with twofold symmetry and triabodies comprised three Fv modules with threefold symmetry. No attempt was made to model conformational changes in the Fv domains. The figures were prepared using Molscript (Kraulis, 1991).

Construction and expression of NC10 scFv with linkers of one, two, three and four glycine residues

The starting template for construction of the scFvs was the 'zero-linked' NC10 scFv–0 gene construct in the vector pPOW (Kortt *et al.*, 1997), in which the 3′ end of the V_H sequence (codon for Ser^{H112}) is linked directly to the 5′ end of the V_L sequence (codon for Asp^{L1}).

Four sets of complementary nucleotides (Table I) were used to mutate pPOW NC10 scFv–0, inserting codons for one, two, three and four glycine residues between codons for Ser^{H112} and Asp^{L1} using QuikChange site-directed mutagenesis (Stratagene Cloning Systems, La Jolla, CA), thus creating scFv–1, scFv–2, scFv–3 and scFv–4. In each reaction, 15 ng of pPOW NC10 scFv–0 plasmid DNA were subjected to PCR in a 50 µl reaction volume containing 5 µl of reaction buffer, as supplied with the kit, 20 pmol of the purified complementary oligonucleotide primer pairs, 2.5 nmol of each dNTP and 2.5 units of *Pfu* DNA polymerase. Thermal cycling conditions were as follows: (95°C, 30 s) 1 cycle; (95°C, 30 s; 55°C, 1 min; 68°C, 12 min) 18 cycles. A 1 µl volume of *Dpn*I restriction enzyme (20 U/µl) was then added to each sample and incubated at 37°C for 90 min. *E.coli* XL1-Blue cells (Stratagene, 1×10^9 cfu/µg) were then transformed by electroporation with 2 µl of each reaction mix at 1.8 kV in 0.1 cm pathlength cuvettes (BioRad Laboratories). After the addition of 500 µl of SOC medium, aliquots of the cells were plated on to YT/amp_{200} and incubated overnight at 30°C. Ten individual colonies were selected from each transformation, plasmid DNA was prepared and sequenced across the region targeted for mutation, using Sequenase version 2.0 (Amersham Life Sciences) with the oligonucleotide primer TACATGCAGCTCAGCAGCCTGAC, a sequence in the V_H region 5′ to the mutation site.

Clones containing the correct sequence insertion were subjected to expression in 5 ml of YT/amp_{200} as described by Malby *et al.* (1993). A small culture sample (10 µl) was probed by Western blot analysis using anti-FLAG M2 antibody, which reacts with the C-terminal epitope tag (FLAG) present in the constructs (Hopp *et al.*, 1988). A positive reaction indicated the presence of a full length, in-frame clone. One clone was selected from each mutation strategy for further scale-up.

Each pPOW-scFv construct was expressed in 500 ml of 2YT/amp_{200} as described by Malby *et al.* (1993), using *E.coli* strain HB2151 as host. The scFv protein was located in the periplasm as insoluble aggregates associated with the membrane fraction, as found previously for NC10 scFv–15, –5 and –0 (Malby *et al.*, 1993; Kortt *et al.*, 1997). The scFv protein was purified as described by Kortt *et al.* (1997). Cells were sonicated in 100 ml of phosphate-buffered saline, pH 7.4 (PBS), centrifuged and the cell pellet was extracted with 50 ml of 6 M guanidinium hydrochloride, 0.1 M Tris–HCl, 5 mM EDTA, pH 8.0. The extract was dialysed against PBS and insoluble material removed by centrifugation. The soluble fraction (supernatant) was then passed over an affinity column of immobilized anti-FLAG M2 antibody (Eastman Kodak, New Haven, CT) and bound protein was eluted with Gentle Elution Buffer (Pierce Chemical, Rockford, IL). The eluted protein was concentrated and dialysed against PBS, pH 7.4, containing 0.02% (w/v) sodium azide and stored frozen.

Estimation of molecular mass of NC10 short-linkered constructs

The relative molecular mass of each affinity-purified NC10 short-linkered scFv was compared by size-exclusion gel chromatography on a Superose 12 HR 10/30 column (Pharmacia) in PBS, pH 7.4, calibrated with BioRad Gel Filtration Standard proteins (thyroglobulin, M_r 670 kDa; bovine γ-globulin, M_r 158 kDa; chicken ovalbumin, M_r 44 kDa; equine myoblobin, M_r 17 kDa; vitamin B-12 M_r 1.35 kDa). The flow-rate was 0.5 ml/min and the absorbance of the effluent stream was monitored at both 214 and 280 nm. Peak elution times were also compared with those of NC10 scFv–0 trimer and NC10 scFv–5 dimer from previous runs on the same column.

Formation of complexes with 3–2 G12 anti idiotype Fab′

Fab′ fragments of the NC10 anti-idiotype antibody 3–2 G12 were prepared as described by Kortt *et al.* (1997). Purified NC10 scFv–2 triabody and scFv–3 diabody from Superose 12 gel chromatography were mixed with a small excess of 3–2G12 anti-idiotype Fab′ in approximately 1:3 and 1:2 molar ratios, respectively, as described by Kortt *et al.* (1997). The

complexes were separated from excess Fab′ by size-exclusion chromatography on Superose 6 (HR 10/30) in PBS, pH 7.4, with a flow-rate of 0.5 ml/min. The column had previously been calibrated with preformed scFv–5/Fab′ and scFv–0/Fab′ complexes and uncomplexed scFv–2, scFv–3 and 3–2G12 Fab′.

Electron microscopy

Complexes of scFv–2 and scFv–3 with 3–2 G12 Fab′ were prepared for electron microscope imaging and data analysis as described previously (Lawrence *et al.*, 1998). Complexes of NC10 Fab′ and influenza neuraminidase (soluble tetrameric extracellular domain) were prepared according to Malby *et al.* (1994) and imaged as described previously (Lawrence *et al.*, 1998). The complexes were diluted in PBS–1% glutaraldehyde to concentrations of the order of 0.01–0.03 mg/ml, applied to glow-discharged carbon-coated 400-mesh gold grids and stained with 2% potassium phosphotungstate at pH 6.0. Micrographs were recorded at 60 kV in a JEOL 100B transmission electron microscope at a magnification of 100 000×. Magnification was calibrated as described by Tulloch *et al.* (1986) by recording the size of influenza virus neuraminidase (100 Å) in single molecule images when complexed with NC10 Fab′.

Particle dimensions (arm lengths and internal inter-arm angles) were measured from the coordinates of digitized micrographs of 211 diabodies and 128 triabodies as described by Lawrence *et al.* (1998). Interpretation of the observed angles was based upon considerations presented by Lawrence *et al.* (1998). In particular, diabody projections were selected only if similar Fab arm lengths were observed in the molecule to ensure that the observed internal inter-arm angle represented the angle between Fab arms lying flat on the support film. Diabody projections with Fab arms of unequal length were disregarded since the Fab arms could be tilted out of the plane of the film or could arise from staining artifacts, radiation damage or dehydration of the protein.

Results

Molecular modelling of shorter linker scFvs

The molecular models (Cα traces) depicted in Figure 1 represent single conformations of NC10 scFv diabodies and triabodies in which the linker length between the C-terminus of V_H and N-terminus of V_L domains is reduced from four glycine residues to zero. All the conformations in Figure 1 are plausible and there are no steric clashes between side-chain atoms. Dimers with four and three residue linkers are shown in two different orientations, each rotated 90° to the other, with the scFv–4 diabody depicted in Figure 1a and c and the scFv–3 diabody depicted in Figure 1b and d. The orientation in Figure 1a and b corresponds to the that reported in Perisic *et al.* (1994). In the diabodies (Figure 1a–d), the regions of the Fv modules in closest proximity to each other are between the 'backs' or C-terminal surface of the V_H domains across the twofold symmetry axis (i.e. between loops distal to the CDRs). As the scFv linker is reduced from four residues to three, the range of orientations between the two Fvs that were sterically allowed was progressively restricted by contacts between the V_H domains (compare Figure 1c for scFv–4 with Figure 1d for scFv–3). When the scFv linker was further reduced to two residues (scFv–2), the diabody could not be modelled as there were no viable orientations of the Fv domains free from side-chain steric clashes. Instead, three scFv–2 molecules could be modelled into a triabody association with threefold symmetry (Figure 1e) with a range of allowed orientations between Fv modules. The regions of the Fv modules in closest proximity in triabodies are the V_H domains, between loops distal to the CDRs. Clashes between V_H domain loops restricted the range of allowed Fv orientations when the linker was reduced from two residues to direct ligation of V_H and V_L domains together as an scFv–0 'zero linker' molecule (Figure 1f). The scFv constructions described below provide protein chemical data on the state of these diabodies and triabodies, both in solution and when complexed with anti-idiotype Fab′.

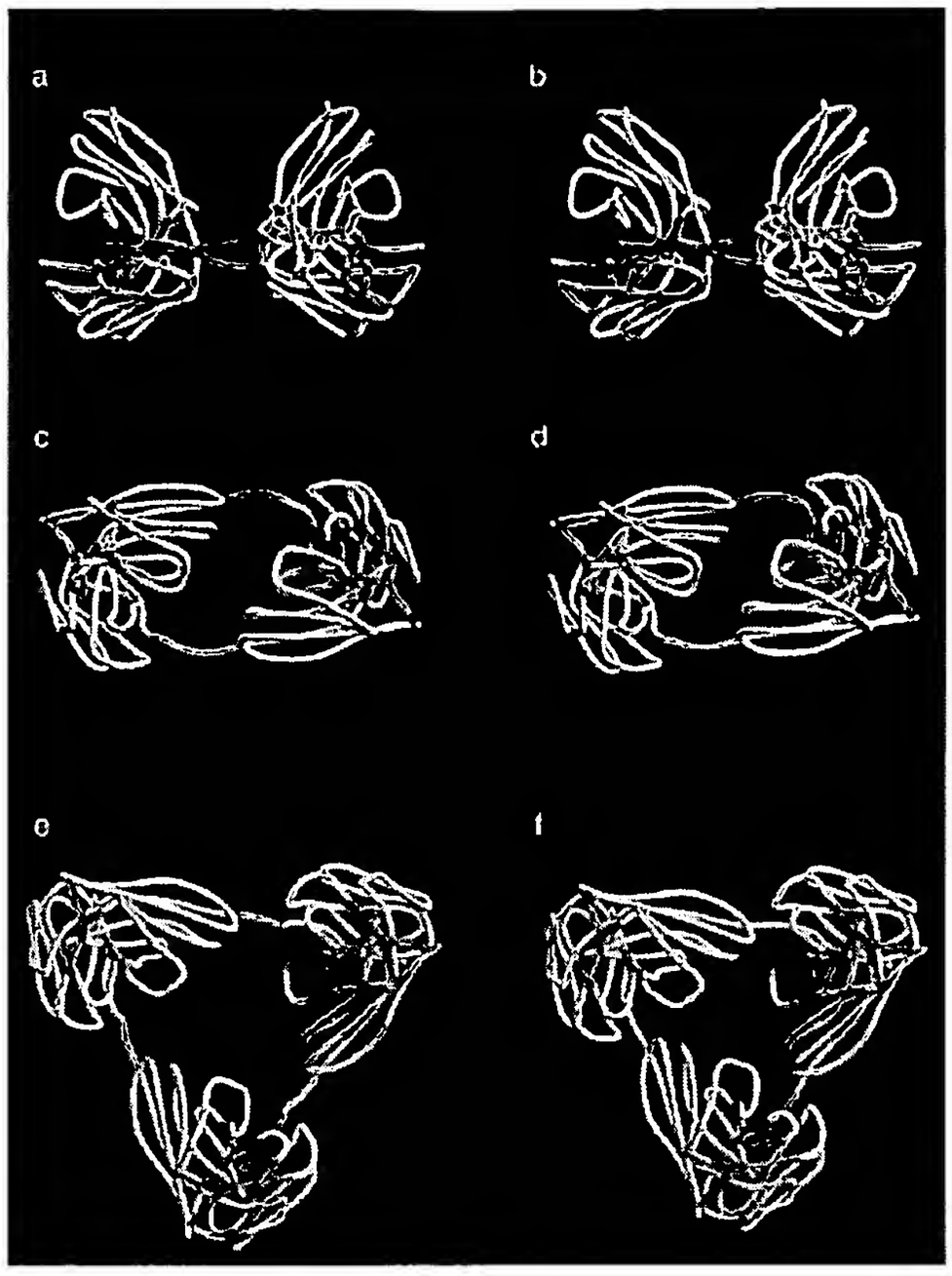

Fig. 1. Molecular models (Cα traces) of scFv dimers and trimers. ScFv–4 (four residue linker) and scFv–3 (three residue linker) are shown in a and b, respectively, the orientation corresponding to that for scFv–5 described by Perisic *et al.* (1994). ScFv–4 and scFv–3 are shown with a 90° rotation (c and d, respectively) to highlight the proximity between V_H domains. ScFv–2 and scFv–0 (trimers with two and zero residue linkers; e and f, respectively) are shown with the threefold symmetry axis perpendicular to the page similar to the model described by Kortt *et al.* (1997). The figures were prepared using Molscript (Kraulis, 1991). The Cα backbone for the different scFv molecules is coloured white and pink for the two scFv molecules in the dimers and the third scFv in the trimer is coloured purple. CDR loops are highlighted in the V_H and V_L domains in blue and yellow, respectively, and the linker between V_H and V_L domains is coloured green.

Construction of NC10 scFvs with linkers of one to four glycine residues

Four NC10 scFvs with short linkers, scFv–1, scFv–2, scFv–3 and scFv–4, were constructed in the pPOW expression vector as described in Materials and methods. The DNA sequence of each scFv construct was confirmed in both orientations to ensure that the correct sequence had been inserted between the V_H and V_L domains (data not shown). Each scFv construct was assessed by expression and Western blot analysis on SDS–PAGE using mouse anti-FLAG M2 antibody and shown to encode an scFv of the expected size including the C-terminal

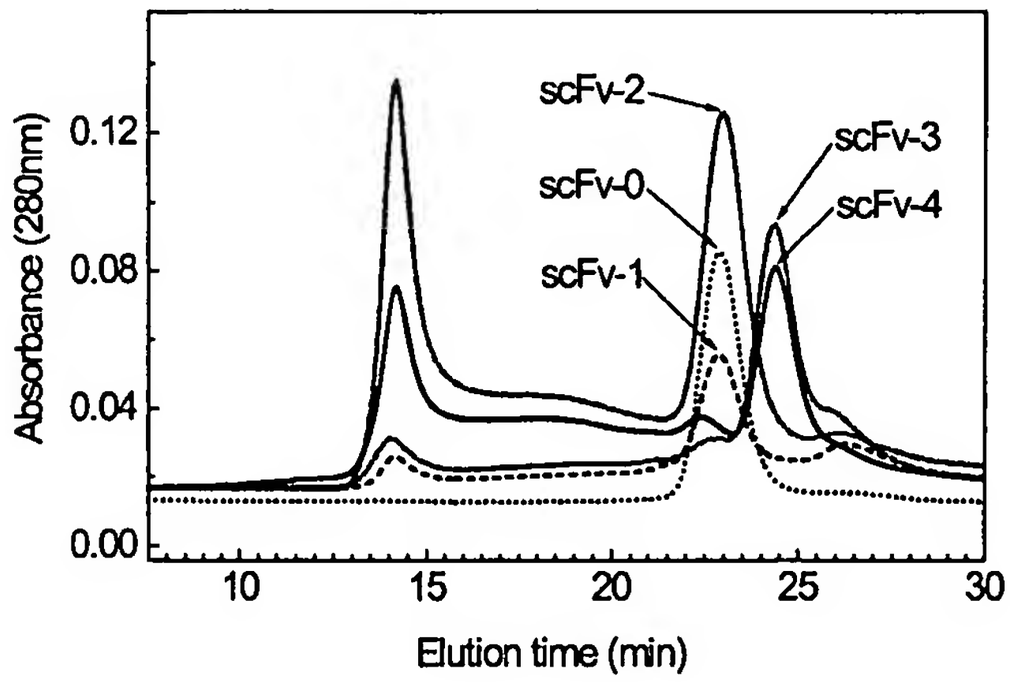

Fig. 2. Size-exclusion chromatography on a calibrated Superose 12 HR 10/30 column of affinity purified NC10 scFvs. Superimposed on the same scale are runs for scFv–1 (- - -), scFv–2 (-·-·-), scFv–3 (-···-···-), scFv–4 (——) and scFv–0 triabody (·····). The major peak of scFv–1 and the scFv–2 peak eluted at a similar time to the scFv–0 triabody (~23 min, ~70 kDa) whereas scFv–3 and scFv–4 eluted at a time (~24.5 min, ~54 kDa) similar to the scFv–5 diabody (not shown). The column was equilibrated in PBS, pH 7.4, and the flow-rate was 0.5 ml/min.

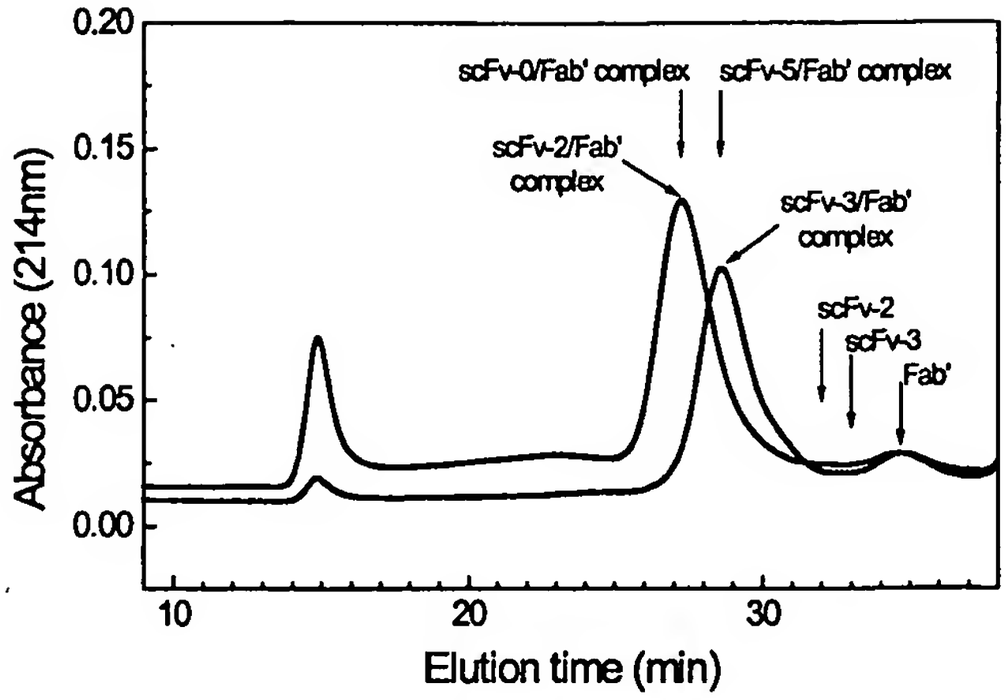

Fig. 3. Size-exclusion chromatography on Superose 6 of complexes 0f scFv–2 and 3–2G12 Fab and of scFv–3 and 3–2G12 Fab′ formed as described in Materials and methods. The column was equilibrated in PBS, pH 7.4, and the flow-rate was 0.5 ml/min. Elution times of uncomplexed scFv–2 triabody, scFv–3 diabody, 3–2G12 Fab′, scFv–0–Fab′ complex and scFv–5–Fab′ complex are indicated by arrows. 3–2G12 Fab′ shows an anomalous elution time, similar to an scFv monomer (A.A.Kortt, unpublished observations).

FLAG tag epitope (~27 kDa; data not shown). The scFv product was recovered in the supernatant fraction following solubilization of the bacterial pellet in 6 M guanidine hydrochloride, dialysis against PBS and centrifugation to remove insoluble material. The scFvs were affinity purified using an anti-FLAG M2 antibody column (Kortt *et al.*, 1997).

Molecular mass of short-linkered NC10 scFvs

Samples of affinity-purified scFv–1, scFv–2, scFv–3 and scFv–4 were shown by SDS–PAGE and Western blot analysis to comprise predominantly scFv of the expected molecular mass (~27 kDa). The samples were individually subjected to analysis by size-exclusion gel chromatography on a calibrated Superose 12 column (Figure 2). The major peaks of each of the scFvs showed the following distribution. Peaks of scFv–1 and scFv–2 eluted at the same time as the zero-linked scFv–0 (~23 min; $M_r \approx 70\,000$ Da), previously shown to be a trivalent trimer (Kortt *et al.*, 1997). The major peaks for scFv–3 and scFv–4 showed elution times identical with that found previously for NC10 scFv–5 (~24.5 min; $M_r \approx 52\,000$ Da), previously shown to be a bivalent dimer (Kortt *et al.*, 1997). The results clearly demonstrate a precise delineation into two groups; scFv–1 and scFv–2 formed trimers similar to scFv–0 whereas scFv–3 and scFv–4 formed dimers similar to scFv–5. The small peak eluting at ~23 min in the scFv–3 and scFv–4 profiles could indicate the presence of small amounts of trimer. There was insufficient material to investigate this further. The small shoulder eluting at 26.5 min is possibly Fv monomer, produced by proteolytic cleavage as observed by Kortt *et al.* (1997). Compared with the size of the major peaks, these minor components represented between 10 and 15% of the recovered recombinant product.

The high molecular mass material at the column void volume (~14 min) had previously been shown (for scFv–0; Kortt *et al.*, 1997) to contain soluble aggregates of scFv.

Formation of complexes with 3–2 G12 anti idiotype Fab′ fragments

3–2G12 is an anti-idiotype monoclonal antibody which competes with influenza virus neuraminidase for binding to NC10. Complexes were formed by mixing scFv–2 or scFv–3 isolated by gel filtration in a 1:3 or 1:2 molar ratio, respectively, with the Fab′ fragment of 3–2G12, an anti-idiotype antibody. The Fab′ was kept in slight excess of these ratios to ensure complete decoration of all antigen binding sites present. The complexes were separated from excess Fab′ on a previously calibrated Superose 6 column (Figure 3). Elution times of uncomplexed scFv–2, scFv–3 and the complexes of scFv–0/Fab′ and scFv–5/Fab′ (Kortt *et al.*, 1997) are indicated by arrows. The complexes eluted as single peaks, well separated from unbound Fab′. There was no evidence of any uncomplexed scFv–2 or scFv–3 in either separation. The complex formed between scFv–2 and 3–2G12 Fab eluted as a single peak at 27.4 ± 0.15 min averaged from several chromatographic analyses. The elution time is identical with that observed for the scFv–0/Fab′ complex, which was found to have a molecular mass of ~235 000 Da (Kortt *et al.*, 1997) and is consistent with the mass of a complex of three Fab′ molecules ($M_r \approx 52\,000$ Da each) and one scFv–2 trimer ($M_r \approx 78\,500$ Da). The scFv-3 complexed with 3–2G12 Fab′ eluted at 28.6 ± 0.1 min, averaged from several chromatographic analyses. The elution time is identical with that observed for the scFv–5–Fab′ complex, shown previously to have a molecular mass of ~156 000 Da (Kortt *et al.*, 1997), which is consistent with a complex comprising two 3–2G12 Fab′ molecules and one scFv–3 dimer ($M_r \approx 52\,000$ Da).

The estimated molecular mass of scFv–anti-idiotype Fab′ complexes is consistent with the prediction that scFv–3 diabodies are bivalent and bind two Fab′ fragments and that scFv–2 triabodies are trivalent and bind three Fab′ fragments.

Electron microscopy of NC10 scFvs complexed with 3–2G12 Fab′ fragments

The complex of scFv–3 and 3–2G12 Fab′ was isolated by gel filtration on Superose 6 (Figure 3) and imaged by electron microscopy. Images of the complex appeared predominantly as boomerang-shaped projections (each arm ~100–110 Å in length) with the internal angle between the two arms ranging from 60 to 180° (Figure 4a). These images were similar to

Fig. 4. Electron micrographs of scFv–Fab′ complexes stained with 2% potassium phosphotungstate. Magnification bar is 50 nm. (a) Montage showing two fields of scFv–3 complexed with 3–2G12 Fab. While the majority of scFv-3 are diabodies, one triabody (circled) is included. (b) Field of scFv–2 complexes with 3–2G12 Fab showing both V- and Y-shaped projections with variable angles between Fab arms. (c) Field showing nine projections of complexes of NC10 Fab with influenza neuraminidase (soluble tetrameric extracellular domain). The nine projections appear X-shaped with the four NC10 Fab arms extending radially from the core neuraminidase tetramer.

those observed previously for scFv–5 diabodies complexed with 3–2G12 Fab′, which showed similar projections consistent with two Fab molecules extending outwards from the antigen binding sites (Lawrence *et al.*, 1998). Despite the potential 'elbow' flexibility between Fv and C modules in the Fab′, each Fab′ arm appeared as a relatively rigid, linear molecular rod (Tulloch *et al.*, 1986; Lawrence *et al.*, 1998). Relatively rigid Fab arms were also observed in free 3–2G12 anti-idiotype Fabs under the same imaging conditions (data not shown) and also in complexes of NC10 Fab with influenza neuraminidase tetramers (Figure 4c). In the neuraminidase complexes, the four NC10 Fab arms extend radially from the central neuraminidase core, consistent with the high-resolution crystal structure (Malby *et al.*, 1998).

Table II. Distribution of inter-arm angles for NC10 scFv–3 and scFv–5 diabodies

Diabody		Inter-arm angle (°)
scFv–3	Minimum	60
	Maximum	179
	Mean ± SD	116.5 ± 20.3
scFv–5	Minimum	60
	Maximum	178
	Mean ± SD	122.4 ± 24.4

In the scFv–3 diabody complexes with 3–2G12 Fab′ (Figure 4a) the arm length and internal angles of the boomerang-shaped projections were measured from digitized micrographs. The mean angle for scFv–3 (116°) was similar to that for scFv–5 (122°), with an approximately normal distribution of angles about the mean (Table II). There were also present some Y-shaped projections, interpreted as tripod-shaped objects previously observed as characteristic of trimers complexed with 3–2G12 Fab′. Based on inspection of over 200 scFv–3 complex images on two separate electron micrographs, these accounted for <3% of the total.

ScFv–2 complexed with 3–2G12 Fab′ appeared as both Y- and V-shaped projections (Figure 4b), similar to the scFv–0 triabody–Fab′ complex imaged previously (Lawrence *et al.*, 1998). The Y-shaped projections (Figure 4b) were interpreted as tripod-shaped objects (viewed from above), with all three legs (i.e. the distal ends of the three Fab molecules) in contact with the carbon film. The Y-shaped projections were unlikely to be planar as invariably at least one of the Fab legs appeared foreshortened. The length of the legs appears to be ~90–100 Å, consistent with a non-planar projection of a tripod.

The V-shaped projections (Figure 4b) were interpreted as tripod-shaped objects lying on their sides on the carbon film, with one Fab′ leg extending upward and partially out of the stain. The vertical Fab′ leg increased the projected density at the base of the V-shaped structures (Figure 4b) and produced a very different image to the boomerang-shaped diabody complexes which frequently had decreased stain at the centre (Figure 4a). This is consistent with the expected models (Figure 1). The interpretation of tripods lying on their side is also consistent with the appearance of a few projections with all three Fab′ legs pointing in the same direction.

Measurement of inter-arm angles for the Y-shaped projections of scFv–2 and scFv–0 triabody complexes showed similar distributions of all angles measured (Table III). The three Fab legs were separated by angles of mean 88, 121 and 149° (Table III) with a significant difference between the smallest and largest angles. However, the range of angles was such that for ~10% of particles the legs were evenly spaced with angles all 120° (±5°), consistent with the model of threefold symmetry (Figure 1).

Discussion

A series of short-linkered scFvs of the anti-neuraminidase antibody NC10 have been produced in which the linker length between V_H and V_L domains was sequentially decreased from four glycine residues (scFv–4) to one glycine residue (scFv–1). Previous studies have established that NC10 scFv–5 (five

Table III. Distribution of inter-arm angles for NC10 scFv–2 and scFv–0 triabodies[a]

Angle	scFv–2 triabody			scFv–0 triabody		
	a	*b*	*c*	*a*	*b*	*c*
Maximum (°)	117	146	178	116	146	180
Minimum (°)	48	85	123	40	94	124
Mean (°)	88	121	149	80	122	150
SD (°)	15.5	10.6	13.4	19.0	17.5	18.8

[a]The three angles around the legs of the Y-shaped tripod are termed *a*, *b* and *c*, where *a* is the smallest angle and c is the largest. For a single image, $a + b + c = 360°$.

residue–Gly_4Ser linker) exists predominantly as a dimer (Kortt *et al.*, 1997) and that NC10 scFv–0 (direct ligation of V_H–V_L with no linker) exists predominantly as a trimer (Kortt *et al.*, 1997). Based upon size-exclusion chromatographic profiles on Superose 12 and electron microscopic imaging of complexes of scFv and anti-idiotype Fab′, we conclude that the scFv–3 and scFv–4 predominantly self-associate to form stable dimers, whereas scFv–1 and scFv–2 predominantly form stable trimers. These data establish that if the length of the linker between the V_H and V_L domains of the scFv of NC10 is shortened to less than three residues (glycines in this case), the association state of the scFv shifts from dimer to trimer. Therefore, the transition from dimer to trimer, as the linker length was reduced from three glycine residues to two, is very clear for NC10 scFv (V_H/V_L).

A precise definition of the N- and C-terminal residues of V_L and V_H is required when considering short-linkered scFvs in which the configuration is V_H–linker–V_L. In the case of NC10, the terminal residues were selected after examination of the 2.4 Å resolution structure from X-ray diffraction analysis of the NC10 Fab-NA complexes (Malby *et al.*, 1994, 1998). Ser^{H112} was selected as the V_H C-terminus because it was not in any direct hydrogen-bonded contact with other V_H domain residues, whereas Thr^{H110} and Val^{H111} contacted framework residues Val^{H12} and Ser^{H87}, respectively. The V_L N-terminal residues were defined as Asp–Ile–Glu in which side-chain atoms in Asp^{L1} and Ile^{L2} contacted framework residues Glu^{L3} and Thr^{L97}. Therefore, V_H Ser^{H112} was joined to V_L Asp^{L1} in the gene constructions encoding NC10 scFv–15 using linkers of 15 residues $(Gly_4Ser)_3$ (Malby *et al.*, 1993, Kortt *et al.*, 1994) and scFv–5 diabodies with five residue linkers (Gly_4Ser) (Kortt *et al.*, 1997). Likewise, V_H Ser^{H112} was joined directly to V_L Asp^{L1} for the gene encoding scFv–0 (triabody; Kortt *et al.*, 1997). In the refined crystal structures of NC10 scFv–15 and NC10 scFv–5 complexed with the antigen neuraminidase, the electron density is poor for both the linker region and the V_H C-terminal residue Ser^{H112} (Malby *et al.*, 1998), indicating that there is some flexibility in the designated V_H C-terminus.

Other reported gene constructions encoding diabodies (five residue linkers) or triabodies (directly linked V_H to V_L cassettes) have included an additional residue, Ser^{H113}, as the C-terminus of the V_H domain (Holliger *et al.*, 1993; Perisic *et al.*, 1994; Holliger *et al.*, 1996, Iliades *et al.*, 1997; Pei *et al.*, 1997). Our analysis based upon the NC10 structure predicts that the additional residue is part of the flexible linker, not part of the V-domain framework.

Valency of the NC10 scFv–3 diabody and scFv–2 triabody was determined by forming complexes with anti-idiotype 3–2G12 Fab′ which were purified and subjected to two independent analytical methods. First, the molecular mass was assessed using size-exclusion columns that had been calibrated with preformed complexes of scFv-5 diabodies and scFv-0 triabodies with 3–2G12 Fab′. The mass of scFv-3 complexes was identical with that of scFv-5 diabody complexes (~156 kDa) and consistent with one diabody binding two Fab′ molecules. The mass of scFv-2 complexes was identical with that of scFv-0 triabody complexes (~235 kDa) and was consistent with one triabody binding three Fab molecules. The complexes were stable and could be stored either at 4°C for several weeks or frozen and thawed as required. Based on the rate of complex formation and stability, the functional affinity (avidity) of the diabodies and triabodies is similar to that reported previously for the scFv–5 diabody and scFv–0 triabody (Kortt *et al.*, 1997).

In the second method of analysis, electron microscopy of scFv–3 diabodies complexed with 3G12 Fab revealed boomerang-shaped projections which are interpreted as two Fab arms extending from the antigen binding sites. These images confirmed our previous observations with scFv–5 (Lawrence *et al.*, 1998) that diabodies are unlikely to form the rigid 'back-to-back' structure revealed in the scFv crystal structure for the NC10 scFv-15 (monomer) complexed to NA (Kortt *et al.*, 1994). Analysis of the projections after densitometry revealed a similar angular distribution between Fab arms in the boomerang-shaped projections for scFv–3 compared with scFv–5 (mean 116° and 122°, respectively; Table II). This indicates that there is considerable flexibility in the linker region joining the two Fv modules in an orientation consistent with a 'twisted' or bent diabody structure similar to that modelled by Holliger *et al.* (1993, 1996) (see Figure 1) and confirmed in a crystal structure of a five-residue linker diabody (Perisic *et al.*, 1994). However, in modelling studies of the L5MK16 diabody (Holliger *et al.*, 1996) the predicted angle between the Fv domains was >122° and ranged from 138 to 166°. A few Y-shaped projections (~3%) in the scFv–3 diabody complexes could be interpreted as triabody complexes similar to scFv–0 (Lawrence *et al.*, 1998). No trimers were observed in preparations of NC10 scFv–5/Fab′ complexes (Lawrence *et al.*, 1998). This implies that, unlike scFv–5, the scFv–3 molecules have a weak propensity to associate into trimers rather than dimers. The size-exclusion chromatographic profiles are consistent with a small percentage of trimers eluting before the dimer peak for scFv–3. How do scFv-3 triabodies form? One possible interpretation is that two scFv–3 molecules initially combine by association of a single V_H to a single V_L, allowing combination with a third scFv–3 molecule through the uncomplexed V_H or V_L domains only if the correct diabody pairing is not immediately formed. Formation of trimers with scFv–3 was observed to be more prevalent than with scFv–5 and presumably the shorter three residue linker restricts V-domain flexibility and increases the time required for complete diabody association.

NC10 scFv–2 forms exclusively triabodies and electron micrograph images of complexes with three 3–2G12 Fab molecules showed Y-shaped projections similar to previous images of scFv–0 complexes (Lawrence *et al.*, 1998). Y-shaped projections are characteristic of a trivalent molecule binding three Fab molecules. These images are consistent with the interpretation that the three Fab′ projections of the trimer–Fab′ complex are not coplanar, but are angled together in one direction like the legs of a tripod, consistent with the triabody

model of NC10 scFv–0 (Figure 1). However, triabodies are obviously flexible molecules since only 10% of the projections have three equal inter-arm angles of 120° (±5°), consistent with the rigid model of threefold symmetry (Figure 1e and f). The observed distribution of angles between Fab arms in the scFv–2 triabody complexes ranged around three angles of mean 88, 121 and 149°. These results did not differ significantly from those observed for scFv–0/Fab′ triabody complexes (Lawrence *et al.*, 1998). The molecular models of scFv–2 and scFv–0 triabodies show that the contacts between Fv modules are minimal and would allow considerable flexibility without steric constraints even with scFv–0. V-shaped images exhibited a distinctive feature of increased density at the base of the V that was interpreted as the projection of the third arm of the tripod pointing upwards (Figure 4b). The V-shaped triabody images were clearly different to the boomerang-shaped projections of diabody complexes (Figure 4a), which had distinctly wider inter-arm angles and low density at the central bend of the boomerang. No projections with the appearance of diabody complexes were observed in the triabody complex images. Since the triabody–Fab complexes are stable in solution and all three antigen combining sites have equivalent affinity (Iliades *et al.*, 1997; Kortt *et al.*, 1997), it is unlikely that one Fab leg has dissociated from the complex to form the V-shaped projections. Also, imaging of purified triabody preparations did not reveal free Fab′ molecules in the proportion that would be expected: in fact, very few free Fabs were observed.

Flexibility and binding affinity in diabodies and triabodies

It is tempting to speculate that the trimeric conformation (triabodies) will predominate over dimers (diabodies) for other antibody scFv fragments with a linker length of less than three residues and that the transition between scFv–3 and scFv–2 will be generic. The molecular models depicted in Figure 1 support this interpretation; when the linker is reduced from four residues to three the orientation of Fv modules is restricted by interactions between the V_H domains across the twofold symmetry axis. There are numerous steric clashes when the linker is further reduced to two residues (the scFv–2 could not be built as a diabody in a sterically allowed conformation). However, it is likely that association between Fv modules will depend on both linker flexibility and unique interface interactions. Linker flexibility will be affected by the choice of inserted residues (we used highly flexible glycine residues) and the choice of the C-terminal residue in V_H and N-terminal residue in V_L which might be constrained by interactions to the V-domain framework. Interface interactions across the Fv interface in diabodies and triabodies are almost exclusively between V_H domains when scFvs are in the V_H–V_L orientation (Holliger *et al.*, 1996; Pei *et al.*, 1997; Malby *et al.*, 1998). The detailed interactions across the Fv interfaces will obviously be unique to each scFv and will affect both flexibility and stoichiometry of diabodies and triabodies. We can only hypothesize that NC10 will be representative of other antibody scFv molecules.

The gain in functional affinity through multivalent binding (often termed avidity; Kortt *et al.*, 1997; Pluckthun and Pack, 1997) makes trimeric scFvs attractive for *in vivo* tumour imaging as an alternative reagent to diabodies (Wu *et al.*, 1996; Zhu *et al.*, 1996; Adams *et al.*, 1998; FitzGerald *et al.*, 1998) and multivalent chemical conjugates (Adams *et al.*, 1993; McCartney *et al.*, 1995; Antoniw *et al.*, 1996; Casey *et al.*, 1996). The gain in functional affinity for scFv trimers compared with scFv monomers is apparently due to reduced off-rates which result from both multiple binding and rebinding to the target antigens (Kortt *et al.*, 1997, Pluckthun and Pack, 1997). There is also the added advantage of reduced blood clearance rates for triabodies over diabodies and scFv monomers which is relevant to *in vivo* applications. Multiple binding to surface-bound antigens is dependent on correct alignment and orientation in the Fv modules of diabodies and triabodies. If multiple binding is not sterically possible, particularly for surface-bound antigens, then apparent gains in functional affinity are likely to be small and due only to the effect of increased rebinding, which is dependent on diffusion rates and surface antigen concentration. Antigen orientation also affects the ability of diabodies and triabodies to bind simultaneously to multiple antigens on a cell surface and this factor is particularly important in the design of any therapeutic reagent required to cross-link surface receptors on either the same or adjacent cells (Pack *et al.*, 1995; Pluckthun and Pack, 1997). Indeed, receptor cross-linking and enhanced cell activation have recently been demonstrated using intact Ig molecules conjugated together into flexible dimers (Ghetie *et al.*, 1997). Flexibility in both scFv diabodies and triabodies is evident in our single-molecule imaging studies. The effect of manipulating linker length, sequence and structure on diabody and triabody stability and flexibility can now be analysed using single-molecule imaging of anti-idiotype complexes. The ability of triabodies to cross-link surface receptors is unknown and will obviously depend on flexibility between the Fv modules and the orientation of the antigen binding sites, as well as the structure of the receptor. Furthermore, the construction of tricistronic expression vectors should allow the production of trispecific scFv–0 trimers capable of cross-linking different target antigens, with obvious applications in specific cell recruitment and activation.

Acknowledgements

The authors thank Professors D.Metzger and R.Webster for providing the 3–2G12 and NC10 hybridoma cell lines, Dr D.Hewish for preparing the 3–2G12 antibodies, J.Burns for assistance with protein purification and chromatographic analysis and Drs K.Maxwell and T.Garrett for assistance with modelling using Molscript.

References

Adams,G.P. *et al.* (1993) *Cancer Res.*, **53**, 4026–4034.
Adams,G.P., Schier,R., McCall,A.M., Crawford,R.S., Wolf,E.J., Weiner,L.M. and Marks,J.D. (1998) *Br. J. Cancer*, **77**, 1405–1412.
Antoniw,P. *et al.* (1996) *Br. J. Cancer*, **74**, 513–524.
Atwell,J.L., Pearce,L.A., Lah,M., Gruen,L.C., Kortt,A.A. and Hudson,P.J., (1996) *Mol. Immunol.*, **33**, 1301–1312.
Bird,R.E. *et al.* (1988) *Science*, **242**, 423–426.
Casey,J.L., King,D.J., Chaplin,L.C., Haines,A.M., Pedley,R.B., Mountain,A., Yarranton,G.T. and Begent,R.H. (1996) *Br. J. Cancer*, **74**, 1397–1405.
FitzGerald,K., Holliger,P. and Winter, G. (1997) *Protein Engng*, **10**, 1221–1225.
Ghetie,M.A., Podar,E.M., Ilgen,A., Gordon,B.E., Uhr,J.W. and Vitetta,E.S. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 7509–7514.
Holliger,P., Prospero,T. and Winter,G. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 6444–6448.
Holliger,P., Brissinick,J., Williams,R.L., Thielemans,K. and Winter,G. (1996) *Protein Engng*, **9**, 299–305.
Hopp,T.P., Prickett,K.S., Libby,R.T., March,C.J., Cerretini,D.P., Uradl,D.L. and Conlon,P.J. (1988) *Bio/Technology*, **6**, 1204–1210.
Huston,JS. *et al.* (1988) *Proc. Natl Acad. Sci. USA*, **85**, 5879–5883.
Iliades,P., Kortt,A.A. and Hudson,P.J., (1997) *FEBS Lett.*, **409**, 437–441.
Jones,T.A., Zou,J-Y., Cowan,S.W.,and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **B47**, 110–119.
Kabat,E.A., Wu,T.T., Perry,H.M., Gottensman,K.S. and Foeler,C. (1991) *Sequences of Proteins of Immunological Interest.* US Department of Health

and Human Service, US Public Health Service, National Institutes of Health, Bethesda, MD.
Kortt,A.A. *et al.* (1994) *Eur. J. Biochem.*, **221**, 151–157.
Kortt,A.A. *et al.* (1997) *Protein Engng*, **10**, 423–433.
Kraulis,P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.
Lawrence,L.J., Kortt,A.A., Iliades,P., Tulloch,P.A. and Hudson,P.J. (1998) *FEBS Lett.*, **425**, 479–484.
Malby,R.L. *et al.* (1993) *Proteins: Struct. Funct. Genet.*, **16**, 57–63.
Malby,R.L., Tulip,W.R., Harley,V.R., McKimm-Breschkin,J.L., Laver,W.G., Webster,R.G. and Colman, P.M. (1994) *Structure*, **2**, 733–746.
Malby,R.L., McCoy,A.J., Kortt,A.A., Hudson,P.J. and Colman, P.M. (1998) *J. Mol. Biol.*, **279**, 901–910.
McCartney,J.E. *et al.* (1995) *Protein Engng*, **8**, 301–314.
McGuinness,B.T., Walter,G., FitzGerald,K., Schuler,P., Mahoney,W., Duncan, A.R. and Hoogenboom,H.R. (1996) *Nature Biotechnol.*, **14**, 1149–1154.
Pack,P., Muller,K., Zahn,R. and Pluckthun,A., (1995) *J. Mol. Biol.*, **246**, 28–34.
Pei,X,Y., Holliger,P., Murzin,A.G. and Williams,R.L. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 9637–9642.
Perisic,O., Webb,P.A., Holliger,P., Winter,G. and Williams,R.L., (1994) *Structure*, **2**, 1217–1226.
Pluckthun,A. and Pack,P. (1997) *Immunotechnology*, **3**, 83–105
Tulloch,P.A., Colman,P.M., Davis,P.C., Laver,W.G., Webster,R.G. and Air,G.M. (1986) *J. Mol. Biol.*, **190**, 215–225.
Whitlow,M., Filpula,D., Rollence,M.L., Feng,S.-L. and Woods,J.F. (1994) *Protein Engng*, **7**, 1017–1026.
Wu,A.M., Chen,W., Raubitschek,A., Williams,L.E., Neumaier,M., Fischer,R., Hu,S.-Z., Odom-Maryon,T., Wong,J.Y.C. and Shively,J.E. (1996) *Immunotechnology*, **2**, 21–36.
Zdanov,A., Li,Y., Bundle,D.R., Deng,S.-J., MacKenzie,C.R., Narang,S.A., Young,N.M. and Cygler,M. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 6423–6427.
Zhu,Z., Zapata,G., Shalaby,R., Snedecor,B., Chen,H. and Carter,P. (1996) *Nature Biotechnol.*, **14**, 192–196.

Received October 22, 1998; revised March 26, 1999; accepted 16 April, 1999